

APPROXIMATE VARIANCES FOR TAPERED SPECTRAL ESTIMATES

Michael Amrein and Hans R. Künsch

Seminar für Statistik
ETH Zentrum
CH-8092 Zürich, Switzerland

January 2011

Abstract

We propose an approximation of the asymptotic variance that removes a certain discontinuity in the usual formula for the raw and the smoothed periodogram in case a data taper is used. It is based on an approximation of the covariance of the (tapered) periodogram at two arbitrary frequencies. Exact computations of the variances for a Gaussian white noise and an AR(4) process show that the approximation is more accurate than the usual formula.

Key words: Asymptotic variance, data taper, (smoothed) periodogram.

1 Introduction

Spectral estimation is by now a standard topic in time series analysis, and many excellent books are available, e.g. Percival and Walden (1993) or Bloomfield (2000). The purpose of this short note is to propose an approximation of the asymptotic variance that removes a certain discontinuity in the usual formula for the raw and the smoothed periodogram in case a taper is used. The standard asymptotic variance of the raw periodogram is independent of the taper chosen, see Formulae (222b) and (223c) in Percival and Walden (1993). However, this changes when the raw periodogram is smoothed over frequencies close by. Then a variance inflation factor C_h , see (4), appears which is equal to one if no taper is used and greater than one otherwise, compare Table 248 in Percival and Walden (1993). The reason for this is that tapering introduces correlations between the raw periodogram at different Fourier frequencies. Because of this, the variance reduction due to smoothing is smaller in the case of no tapering. The above variance inflation factor is justified asymptotically when the number of Fourier frequencies that are involved in the smoothing tends to infinity (more slowly than the number of observations, otherwise we would have a bias). Hence,

if only little smoothing is used, then we expect something in between: some increase in the variance, but less than the asymptotic variance inflation factor C_h . We give here a formula, see (5), which is almost as simple as the inflation factor, but which takes the amount of smoothing into account.

2 Notation and preliminaries

Let $\{X_t\}_{t \in \mathbb{Z}}$ be a real-valued stationary process with observation frequency $1/\Delta$, mean $E[X_t] = \mu$, autocovariances $s_\tau := \text{Cov}(X_t, X_{t+\tau})$ and spectral density $S(f)$. We assume that X_1, X_2, \dots, X_N have been observed. The tapered periodogram (called direct spectral estimator in Percival and Walden (1993)) is

$$\hat{S}^{(tp)}(f) := \frac{\Delta}{\sum_{t=1}^N h_t^2} \left| \sum_{t=1}^N h_t (X_t - \tilde{\mu}) e^{-i2\pi i t f \Delta} \right|^2$$

for $f \in [0, 1/(2\Delta)]$. Here the estimator $\tilde{\mu}$ is usually either the arithmetic mean \bar{X} or the weighted average $(\sum_{t=1}^N h_t X_t) / (\sum_{t=1}^N h_t)$. The latter has the property that $\hat{S}^{(tp)}(0) = 0$. Since the choice is irrelevant for the asymptotics, we can use either version. The taper (h_1, \dots, h_N) is chosen to reduce the discontinuities of the observation window at the edges $t = 1$ and $t = N$. Usually, it has the form $h_t = h((2t-1)/(2N))$ with a function h that is independent of the sample size N . A popular choice is the split cosine taper

$$h^p(x) = \begin{cases} \frac{1}{2}(1 - \cos(2\pi x/p)) & 0 \leq x \leq \frac{p}{2} \\ 1 & \frac{p}{2} < x < 1 - \frac{p}{2} \\ \frac{1}{2}(1 - \cos(2\pi(1-x)/p)) & 1 - \frac{p}{2} \leq x \leq 1 \end{cases}. \quad (1)$$

The tapered periodogram has the approximate variance

$$\text{Var}[\hat{S}^{(tp)}(f)] \approx S(f)^2, \quad f \notin \{0, 1/(2\Delta)\} \quad (2)$$

(see e.g. Percival and Walden (1993), Formula (222b)). In particular, it does not converge to zero. Because of this, one usually smoothes the periodogram over a small band of neighboring frequencies. We smooth discretely over an equidistant grid of frequencies. Let $f_{N',k} = k/(N'\Delta)$ ($0 \leq k \leq N'/2$) for an integer N' of order $O(N)$. Then the tapered and smoothed spectral estimate is

$$\hat{S}^{(ts)}(f_{N',k}) = \sum_{j=-M}^M g_j \hat{S}^{(tp)}(f_{N',k-j}),$$

where the g_j 's are weights with the properties $g_j > 0$, $g_j = g_{-j}$ ($-M \leq j \leq M$) and $\sum_{j=-M}^M g_j = 1$. If $k \leq M$, the smoothing includes the value $\hat{S}^{(tp)}(0)$ which is equal or very close to zero if the mean μ is estimated. In this case, we should exclude $j = k$ from the sum.

3 Approximations of the variance of spectral estimators

The usual approximation for the relative variance of $\hat{S}^{(ts)}(f_{N',k})$ is

$$\text{Var} \left(\frac{\hat{S}^{(ts)}(f_{N',k})}{S(f_{N',k})} \right) \approx C_h \frac{N'}{N} \sum_{r=-M}^M g_r^2 \quad (3)$$

for $k \neq 0, N'/2$ where

$$C_h = \frac{\sum_{t=1}^N h_t^4 / N}{(\sum_{t=1}^N h_t^2 / N)^2}. \quad (4)$$

This formula is given in Bloomfield (2000), equation (9.12) on p. 183, and it is implemented in the function “spec.pgram” in the language for statistical computing R (R Development Core Team (2010)). In order to see that it is the same as Formula (248a) in Percival and Walden (1993), one has to go back to the definition of W_m in terms of the weights g_j which is given by the formulae (237c), (238d) and (238e). If we put $M = 0$, (3) is different from (2). The reason for this difference is that (3) is valid in the limit $M \rightarrow \infty$ and $M/N' \rightarrow 0$. But in applications M is often small, e.g. $M = 1$, and one wonders how good the approximation is in such a case.

We propose here as alternative the following approximation for the relative variance

$$\left(\sum_{r=-M}^M g_r^2 + 2 \sum_{l=1}^{2M} \frac{|H_2^{(N)}(f_{N',l})|^2}{H_2^{(N)}(0)^2} \sum_{r=-M}^{M-l} g_r g_{r+l} \right) \quad (5)$$

(again for $k \neq 0, N'/2$) where $H_2^{(N)}(f) = \frac{1}{N} \sum_{t=1}^N h_t^2 e^{-i2\pi t f \Delta}$. In order to compute this expression, we need to compute the convolution of the weights (g_j) and the discrete Fourier transform of the squared taper. The former is usually not a problem since M is substantially smaller than N' . Using the fast Fourier transform, exact computation of the latter is in most cases also possible. If not, then by the Lemma below we can use

$$H_2^{(N)}(f) \approx \int_0^1 h^2(u) e^{-i2\pi N u f \Delta} du e^{-i\pi f \Delta} \frac{\pi f \Delta}{\sin(\pi f \Delta)}.$$

Choosing a simple form for the function h , we can compute the integral on the right exactly. It is obvious that (5) agrees with (2) for $M = 0$. In the next section, we show that it also agrees with (3) for M large.

4 Justification of the approximation

The idea is simple: We just plug in a suitable approximation for the relative covariances of the tapered periodogram values into the exact expression for the

relative variance. $\text{Var} \left(\hat{S}^{(ts)}(f_{N',k})/S(f_{N',k}) \right)$ is equal to

$$\sum_{r=-M}^M \sum_{s=-M}^M g_r g_s \text{Cov} \left(\frac{\hat{S}^{(tp)}(f_{N',k-r})}{S(f_{N',k-r})}, \frac{\hat{S}^{(tp)}(f_{N',k-s})}{S(f_{N',k-s})} \right). \quad (6)$$

The asymptotic behavior of these covariances is well known. Theorem 5.2.8 of Brillinger (1975) shows that, under suitable conditions, we have for frequencies $0 < f \leq g < 1/(2\Delta)$ that

$$\text{Cov} \left(\frac{\hat{S}^{(tp)}(f)}{S(f)}, \frac{\hat{S}^{(tp)}(g)}{S(g)} \right) = \frac{|H_2^{(N)}(f-g)|^2 + |H_2^{(N)}(f+g)|^2}{|H_2^{(N)}(0)|^2} + O(N^{-1}). \quad (7)$$

The statement in Brillinger (1975) is actually asymmetric in f and g since it has $S(f)$ instead of $S(g)$ on the left side in the equation above. Our statement can be proved by the same argument if we assume $S(f) \approx S(g)$ when $|f-g|$ is small. When $|f-g|$ is big, i.e. not of order $O(N^{-1})$, the covariance is of the order $O(N^{-1})$ anyhow. Using the approximation (7) directly would lead to an approximation which depends on k . Having to compute $N'/2$ different approximate variances is usually too complicated. However, the term $|H_2^{(N)}(f+g)|^2$ is small unless $\Delta(f+g)$ is close to zero modulo one. This has been pointed out by Thomson (1977), see also the discussion on p. 230–231 of Percival and Walden (1993). If we omit this term, then we obtain our new approximation (5) by a simple change in the summation indices.

We next give a simple lemma that justifies the omission of the second term in (7). In addition, it also shows how the usual approximation (3) follows from (5).

Lemma 4.1. *If ψ is once continuously differentiable on $[0, 1]$ and ψ' is Lipschitz continuous with constant L , then*

$$\frac{1}{N} \sum_{t=1}^N \psi \left(\frac{2t-1}{2N} \right) e^{-i2\pi\lambda t} = \int_0^1 \psi(u) e^{-i2\pi N\lambda u} du e^{-i\pi\lambda} \frac{\pi\lambda}{\sin(\pi\lambda)} + R$$

where $|R| \leq \text{const.}/N$ uniformly for all $\lambda \in [0, 0.5]$.

Proof. Put $\epsilon = 1/(2N)$. By a Taylor expansion, we obtain for any $x \in [0, 1]$

$$\begin{aligned} \int_{x-\epsilon}^{x+\epsilon} \psi(u) e^{-i2\pi N\lambda u} du &= \psi(x) e^{-i2\pi N\lambda x} \int_{-\epsilon}^{\epsilon} e^{-i2\pi N\lambda u} du \\ &\quad + \psi'(x) e^{-i2\pi N\lambda x} \int_{-\epsilon}^{\epsilon} u e^{-i2\pi N\lambda u} du + R' \end{aligned}$$

where the remainder satisfies $|R'| \leq 2\epsilon^3 L/3 = L/(12N^3)$. Next, observe that

$$\int_{-\epsilon}^{\epsilon} e^{-i2\pi N\lambda u} du = \frac{\sin(\pi\lambda)}{\pi\lambda N}, \quad \int_{-\epsilon}^{\epsilon} u e^{-i2\pi N\lambda u} du = \frac{i}{2\pi\lambda N^2} \left(\cos(\pi\lambda) - \frac{\sin(\pi\lambda)}{\pi\lambda} \right).$$

From this the lemma follows by taking $x = (2t-1)/(2N)$ for $t = 1, \dots, N$ and summing up all terms. \square

If $\psi(0) = \psi(1) = 0$, then by partial integration

$$\left| \int_0^1 \psi(u) e^{-i2\pi N \lambda u} du \right| \leq \frac{\sup |\psi'(x)|}{2\pi \lambda N}.$$

Hence by setting $\psi(u) = h^2(u)$, we obtain

$$H_2^{(N)}(f) \leq \text{const.} (Nf\Delta)^{-1} + \text{const.} N^{-1} \leq \text{const.} (Nf\Delta)^{-1} \quad (8)$$

for $f \leq 1/\Delta$. Therefore the second term in (7) is negligible unless $f + g$ is of the order $O(N^{-1})$.

Finally, we derive the usual variance approximation (3) from (5) as follows. By Parseval's theorem

$$\sum_{l=-N'/2}^{N'/2} \left| H_2^{(N)}(f_{N',l}) \right|^2 = \frac{N'}{N} \frac{1}{N} \sum_{t=1}^N h_t^4.$$

Note that $H_2^{(N)}(0) = 1/N \cdot \sum_{t=1}^N h_t^2$. Because of (8), we have

$$\sum_{l=2M+1}^{N'/2} \left| H_2^{(N)}(f_{N',l}) \right|^2 \leq \sum_{l=2M+1}^{N'/2} \left(\frac{\text{const.} \cdot N'}{Nl\Delta} \right)^2 \rightarrow 0$$

for $M \rightarrow \infty$ and $N' = O(N) \rightarrow \infty$. Thus

$$\sum_{l=-2M}^{2M} \left| H_2^{(N)}(f_{N',l}) \right|^2 - \frac{N'}{N} \frac{1}{N} \sum_{t=1}^N h_t^4 \rightarrow 0,$$

also in the above limit. If the weights g_j change smoothly as a function g of the lag j , i.e., $g_j = g(j/M)$, then for any fixed l

$$\sum_{r=-M}^{M-l} g_r g_{r+l} \sim M \int_{-1}^1 g^2(u) du \sim \sum_{r=-M}^M g_r^2 \quad (M \rightarrow \infty)$$

and the desired result follows by dominated convergence.

5 Comparison with exact relative variances for Gaussian processes

If we assume the process $\{X_t\}_{t \in \mathbb{Z}}$ to be Gaussian, then it holds

$$\begin{aligned} \text{Cov} \left(\frac{\hat{S}^{(tp)}(f)}{S(f)}, \frac{\hat{S}^{(tp)}(g)}{S(g)} \right) &= \frac{1}{S(f)S(g)(\sum_{t=1}^N h_t^2)^2} \\ &\times \left(\left| \sum_{j,k=1}^N h_j h_k s_{j-k} e^{-i2\pi(fj-gk)\Delta} \right|^2 + \left| \sum_{j,k=1}^N h_j h_k s_{j-k} e^{-i2\pi(fj+gk)\Delta} \right|^2 \right), \end{aligned}$$

see p. 326 of Percival and Walden (1993). Plugging this into (6) yields thus an exact expression. Evaluation is of the order $O(N^3)$, thus it is not practical to use it routinely.

We now compare the two approximations to the exact relative variances for a Gaussian white noise $X_t = \epsilon_t$ and the AR(4) process

$$X_t = 2.7607X_{t-1} - 3.8106X_{t-2} + 2.6535X_{t-3} - 0.9238X_{t-4} + \epsilon_t \quad (9)$$

used in Percival and Walden (1993) (see p. 46) where ϵ_t i.i.d $\sim \mathcal{N}(0,1)$. True spectra are shown in Figure 1 in decibel (dB), i.e., the plot displays $10\log_{10}(S(\cdot))$. As we can see, the spectrum of the AR(4) process varies over a

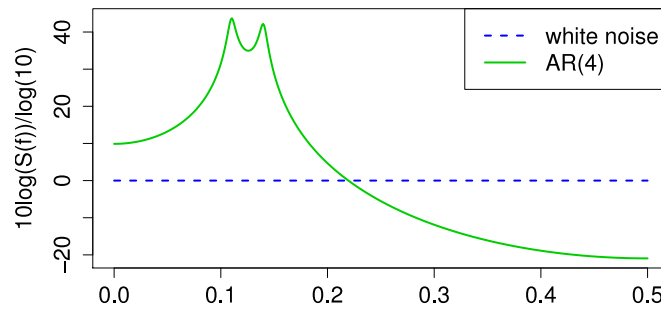


Figure 1: Spectra of the Gaussian white noise and the AR(4) process in (9) in dB.

wide range and exhibits two sharp peaks. Further, we assume the observation frequency $1/\Delta$ to be 1 and $N = 2^{10} = 1024$. We compute the exact relative variance (6) at the frequencies $f_{k,N'}$, $k = 0, \dots, N'/2$, for $N' \in \{N, 2N\}$, the split cosine taper (1) with $p \in \{0.2, 0.5\}$ and weights $g_j = 1/(2M + 1)$, $j = -M, \dots, M$, with $M \in \{0, 1, 2\}$. Comparison to the usual approximation (3) and to the new one (5) is shown in Figure 2. The code in R is available under http://stat.ethz.ch/~kuensch/papers/approximate_variances.R. We see that the new approximation fits the true relative variances clearly better when we smooth over few frequencies, i.e., M is small. Especially in the situations when the data is strongly tapered ($p = 0.5$) or when we use a refined smoothing grid ($N' = 2N$) we recommend to use the new approximation (5).

Acknowledgement

We thank Don Percival and Martin Mächler for helpful comments and suggestions on earlier versions.

References

Bloomfield, P. (2000). *Fourier Analysis of Time Series: An Introduction*. Wiley, NY, 2nd edition.

- Brillinger, D. R. (1975). *Time Series, Data Analysis and Theory*. International Series in Decision Processes. Holt, Rinehart and Winston, Inc., New York.
- Percival, D. B. and Walden, A. T. (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge University Press.
- R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Thomson, D. J. (1977). Spectrum estimation techniques for characterization and development of WT4 waveguide - I. *Bell System Technical Journal*, 56:1769–1815.

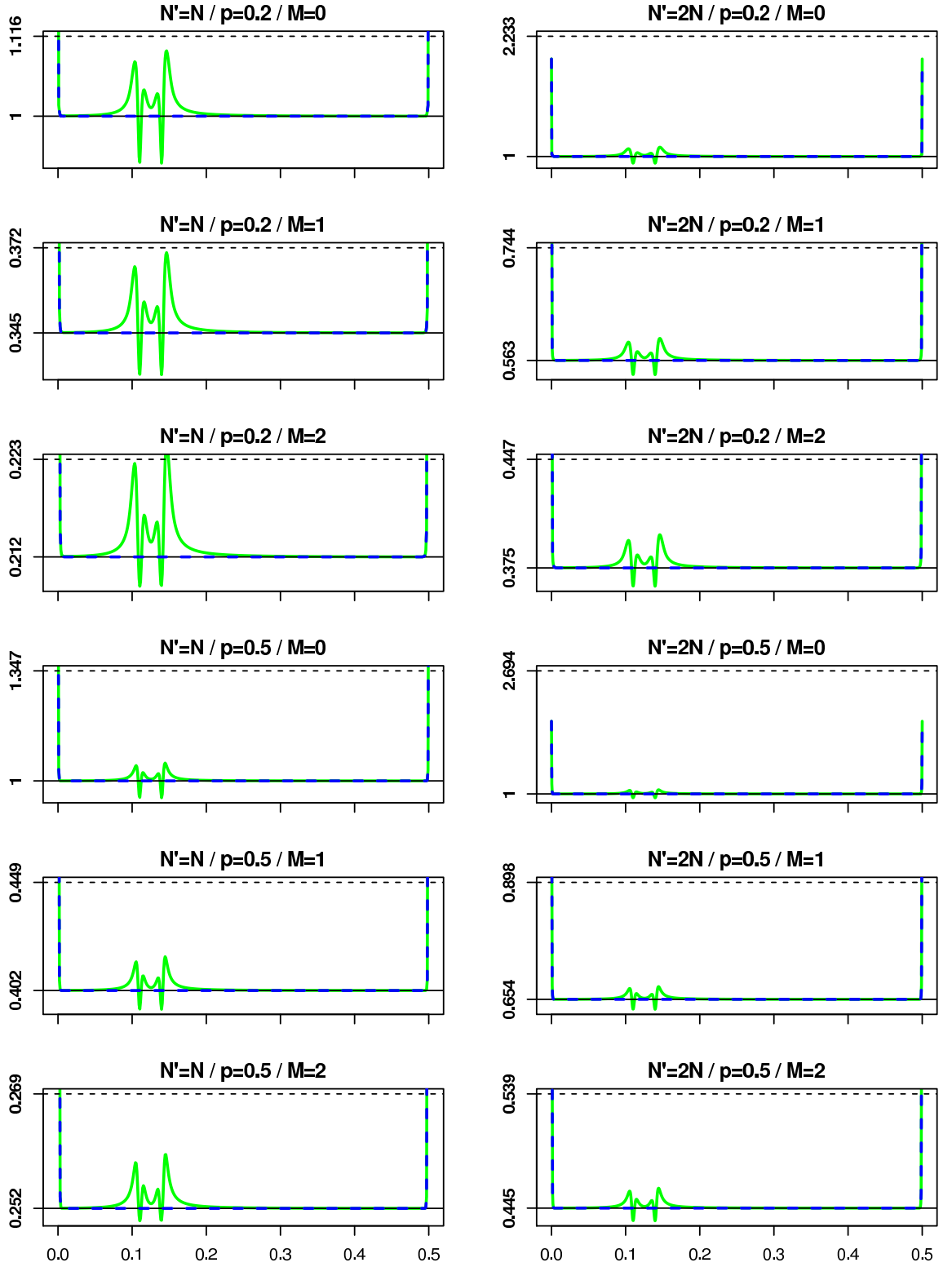


Figure 2: Exact relative variances of the **Gaussian white noise** (**thick-dashed**) and the **AR(4) process** (**thick-solid**) in comparison to the **usual** (**thin-dashed**) and the **new** (**thin-solid**) approximation for different choices of N' , p and M .